The potential of metabolomics for Leishmania research in the post-genomics era

Scheltema, Richard A.; Decuypere, Saskia; T'Kindt, Ruben; Dujardin, Jean-Claude; Coombs, Graham H.; Breitling, Rainer; T'Kindt, Ruben

Published in:
Parasitology

DOI:
10.1017/S0031182009992022

The potential of metabolomics for *Leishmania* research in the post-genomics era

RICHARD A. SCHELTEMA[1]†, SASKIA DECUYPERE[2,3]†, RUBEN T'KINDT[2,3], JEAN-CLAUDE DUJARDIN[2], GRAHAM H. COOMBS[3] *and* RAINER BREITLING[1]*

[1] *Groningen Bioinformatics Centre, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Kerklaan 30, 9751 NN Haren, The Netherlands*
[2] *Department of Parasitology, Unit of Molecular Parasitology, Institute of Tropical Medicine, Antwerp B-2000, Belgium*
[3] *Strathclyde Institute of Pharmacy and Biomedical Sciences, University of Strathclyde, Glasgow G4 0NR, Scotland*

(*Received 27 October 2009; revised 27 November 2009; accepted 27 November 2009; first published online 29 January 2010*)

SUMMARY

The post-genomics era has provided researchers with access to a new generation of tools for the global characterization and understanding of pathogen diversity. This review provides a critical summary of published *Leishmania* post-genomic research efforts to date, and discusses the potential impact of the addition of metabolomics to the post-genomic toolbox. Metabolomics aims at understanding biology by comprehensive metabolite profiling. We present an overview of the design and interpretation of metabolomics experiments in the context of *Leishmania* research. Sample preparation, measurement techniques, and bioinformatics analysis of the generated complex datasets are discussed in detail. To illustrate the concepts and the expected results of metabolomics analyses, we also present an overview of comparative metabolic profiles of drug-sensitive and drug-resistant *Leishmania donovani* clinical isolates.

Key words: Post-genomics research, metabolomics, mass spectrometry, liquid chromatography, *Leishmania*.

THE CHALLENGE OF *LEISHMANIA* DIVERSITY

Leishmania is a group of kinetoplastid protozoa with a highly diverse character. The genus includes more than 20 parasite species infecting humans following transmission by a sand fly (Diptera: Psychodidae: Phlebotominae). All species are pathogenic, but the disease manifestation depends on the infecting species and varies from self-healing skin lesions, to disfiguring mucocutaneous lesions, to lethal visceral disease. Even a single *Leishmania* species can challenge health professionals with enormous clinical diversity; disease severity can vary greatly between different patients, variable response or lack of response to administered treatments is often observed, and post-leishmaniasis syndrome can appear in some, but not all cases. This clinical polymorphism results from the interaction between several factors, including variation in the host, parasite, vector, and environment (Dujardin, 2009). Hardly anything is known about the nature and identity of the intra- and inter-species molecular differences underlying the clinical polymorphism in leishmaniases (Smith *et al*. 2007). Diversity also marks the *Leishmania* life cycle, which consists of two distinct life-forms, promastigotes and amastigotes, adapted to life in a sand fly vector and mammalian host, respectively. The transformation from promastigote to amastigote is essential for the parasite's survival upon transmission from vector to host. How *Leishmania* differentiates between life stages, and which metabolic pathways are essential for survival within each host, are pressing questions requiring an answer in order to develop more effective treatments and alternative control measures against these pathogens in the face of ever increasing resistance to the few available drugs (McConville *et al*. 2007). A detailed molecular characterization of parasite diversity and the underlying biology, including mechanisms of drug resistance, is therefore a pressing need.

Whole-genome sequencing, transcriptomics studies based on whole genome-based DNA arrays, and global proteomic studies have already been carried out in an effort to characterize *Leishmania* diversity. Now, the toolbox of post-genomic research has been complemented by metabolomics, which involves global analysis of the metabolites in a cell (i.e. the metabolome) (Fernie *et al*. 2004; Goodacre *et al*. 2004). Metabolomics is an emerging field in functional genomics that is complementary to the other post-genomic technologies and potentially a powerful tool in its own right to characterize *Leishmania* diversity. This review starts by providing a critical summary of published *Leishmania* post-genomic research efforts that have scrutinized

* Address correspondence to: Rainer Breitling (r.breitling@rug.nl), Tel: +31-50-3638088, Fax: +31-50-3637976.
† Equally contributing authors.

doi:10.1017/S0031182009992022

Leishmania diversity. Secondly, we will discuss in detail the opportunities and challenges of a metabolomics approach specifically for *Leishmania* research, focusing on the critical issues of sample preparation, measurement optimization, and setting up an efficient bioinformatics pipeline for processing the complex data generated. Finally, we will briefly discuss how an integrated approach combining information across all molecular levels holds the greatest potential for understanding parasite biology and successfully developing improved control strategies for the debilitating diseases they cause.

CURRENT CONTRIBUTION OF GENOME, TRANSCRIPTOME AND PROTEOME STUDIES

The whole genomes of *L. major*, *L. infantum* and *L. braziliensis* have been sequenced and comparatively analyzed (Peacock *et al*. 2007). Despite the broad differences in disease pathologies, it appears that the three genomes have a highly conserved gene content, synteny and genome architecture, and only very few genes are species-specific in each genome (5, 26 and 47, respectively). The authors concluded that either (1) a few species-specific genes determine pathogenesis, or (2) the parasite genome plays only a minor role in determining clinical profile, or (3) parasite gene/protein expression levels differ considerably between species, possibly as a consequence of gene dosage in the form of gene copy number variation. The latter hypothesis is supported by comparative studies of *Leishmania* species (benign cutaneous *L. peruviana vs* severe cutaneous *L. braziliensis*) and strains with differential pathogenicity (cutaneous *L. infantum vs* visceral *L. infantum*): in both cases, it was shown that the less pathogenic parasites presented fewer copies of ribosomal DNA (affecting general metabolism) and genes encoding surface antigens (affecting interaction with host cells) compared to the more pathogenic parasites (Victoir *et al*. 1995; Inga *et al*. 1998; Guerbouj *et al*. 2001; Dujardin, 2009). A similar gene-dosage-related phenomenon has repeatedly been described in the context of *Leishmania* drug resistance. Chromosomes containing genes with a functional role in drug resistance were amplified in drug resistant strains, showing polyploidy of the complete chromosome or partial amplification of specific regions as extrachromasomal circular or linear amplicons; this in turn results in up-regulated expression of those amplified genes (Ubeda *et al*. 2008; Ouellette *et al*. 2009; Leprohon *et al*. 2009). Hence, *Leishmania* diversity could very well be encoded by gene copy number variation, either intra- or extrachromosomal. The three-genome project however, was limited with respect to characterization of copy number variation (gene dosage), as the Sanger sequencing technique used in the project starts with cloned genomic DNA fragments that are first amplified *in vivo* (Sturm *et al*. 2008). Meanwhile new sequencing technologies have been developed that can circumvent this limitation (Hall, 2007; Morozova *et al*. 2008). In addition to the relatively low cost and high-throughput capacity, these new technologies offer the advantage of generating sequence data that can be used in a sequence census approach. The sequence census approach involves mapping of all sequence reads to its origin in a reference genome; the number of reads that map to a particular nucleic acid species then correlates with the abundance of that sequence in the cell. This allows comparison of gene dosage at a particular locus in different strains (Morozova *et al*. 2008). In the near future, this new genomic approach is likely to clarify the role gene dosage plays in generating *Leishmania* diversity.

Transcriptomics studies have used *Leishmania*-specific microarrays to assay steady-state levels of mRNA transcripts. Arrays with a genome coverage between 22% and 97·5% have been used to determine differences in expression between different life-cycle stages (Holzer *et al*. 2006; Leifso *et al*. 2007; Saxena *et al*. 2007; Srividya *et al*. 2007; Rochette *et al*. 2008), to determine expression differences associated with drug resistance (Guimond *et al*. 2003; Singh *et al*. 2007; Fadili *et al*. 2009) and clinical polymorphisms (Salotra *et al*. 2006; Singh *et al*. 2007), and to assess expression differences between different species (Rochette *et al*. 2008). However, all these transcriptomics profiling efforts provided few insights compared to experiences with other microbial pathogens. It seems that due to at least two specific factors of *Leishmania* biology, transcriptomics is a strikingly uninformative tool to uncover the molecular background of *Leishmania* diversity. Firstly, *Leishmania* uses polycistronic transcription and very few promoters, and regulation appears to occur at the post-transcriptional level (McDonagh *et al*. 2000; Cohen-Freue *et al*. 2007; Clayton and Shapira, 2007). As a consequence, studies comparing different life stages found that indeed the vast majority of genes are constitutively expressed in all development stages. Only a limited number of genes (2–9%) show preferential stage-specific expression, and those quantitative expression changes are often very small and not always easy to detect reproducibly (Cohen-Freue *et al*. 2007). Secondly, although the *Leishmania* genome sequencing project is complete, more than 60% of the putative protein coding genes do not match annotated sequences and were assigned a hypothetical function (Ivens *et al*. 2005). Hence, the majority of identified differentially expressed genes in any of the above mentioned genome-wide transcriptomics studies turn out to have an unknown biological function. Thus so far, transcriptomics has made a modest contribution in unravelling the cell biological story behind *Leishmania* diversity, as both the function of the identified

proteins and their regulation can not always be adequately assessed.

Proteomic studies have been particular useful for the identification of stage-specific proteins. Given the characteristic expression regulation mechanisms of *Leishmania*, it was not surprising that, compared with the transcriptomics work, proteomic studies reported a higher percentage (12–18%) of the predicted proteome as stage-specifically regulated at the translational or post-translation level (Bente *et al*. 2003; McNicoll *et al*. 2006; Leifso *et al*. 2007; Cohen-Freue *et al*. 2007; Rosenzweig *et al*. 2008). Most of these studies were based on two-dimensional gel electrophoresis; recently the isobaric tagging technique iTRAQ was successfully introduced and proved to be an elegant method for proteomic time-series experiments (Rosenzweig *et al*. 2008). The study profiled 931 proteins over 7 life-cycle differentiation time points, which allowed the inference of a map of the dynamics of most major metabolic pathways during differentiation from promastigotes to amastigotes in *L. donovani*. Proteomics is less rewarding when studying phenotypic or pathologic diversity. Like transcriptomics, proteomic approaches that studied mechanisms of drug resistance (Sharma *et al*. 2003; Drummelsmith *et al*. 2003, 2004; Vergnes *et al*. 2007; Singh *et al*. 2008; Fadili *et al*. 2009) or species diversity (Brobey *et al*. 2006) very often resulted in a list of differentially expressed proteins lacking an annotated function. This again impedes determination of how the different actors relate to each other and often leaves too many open questions.

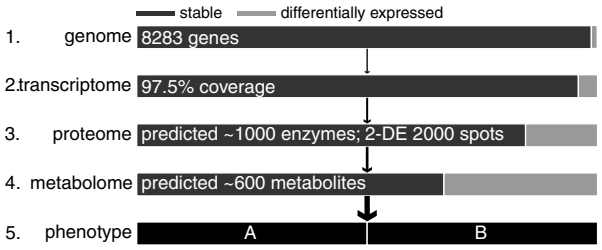

Fig. 1. A rough hypothetical sketch of the progressive variability predicted across molecular levels when moving from *Leishmania* genotype to phenotype. The arrows indicate the predominant chain of causal influence between the levels. (1) The genome is the most stable molecular player of the system, but even at this level ploidy changes and copy number variation might introduce variability in a species and between species (personal communication J. C. Dujardin). (2) Gene expression studies comparing different *Leishmania* life stages or phenotypes, have found that 2–9% of all genes are differentially expressed, while over 90% of genes seem to be constitutively expressed (Cohen-Freue *et al*. 2007). (3) Similar studies based on two-dimensional gel electrophoresis work have found that 5–12% of detected proteins are differentially expressed (Cohen-Freue *et al*. 2007). (4) We predict that the metabolome will be affected to an even greater degree: due to the high connectivity in the metabolic network and multiple metabolites are linked to single enzymes, changes in protein levels are likely to propagate across the metabolome. (5) Ultimately the changes at the molecular level are reflected in the diverse clinical phenotypes of *Leishmania* infections (indicated here as phenotype A and B).

THE POTENTIAL OF METABOLOMICS

In order to fully understand the phenotypic characteristics of an organism, it is essential to complement transcriptome and proteome research with metabolomics investigations. The sequenced genome of *L. major* has already been used to predict the metabolic potential of the organism in two different studies (Chavali *et al*. 2008; Doyle *et al*. 2009). Both studies report around 1000 predicted metabolites, all located in core metabolism, and associated to as many enzyme-catalyzed reactions. As cell-based metabolomics encompasses the different cellular compartments, this number of metabolites can be reduced to around 600, correcting for the metabolites occurring in multiple compartments within the cell.

Changes in protein and transcript levels often function as intermediate steps leading to functionally relevant variation at the metabolite level. It is at the metabolite level that most physiological functions are carried out, making the metabolome the closest correlate to the phenotype (Fig. 1). Metabolite levels are not easily predicted by protein or transcript levels due to the highly non-local control structure of the metabolic network (e.g. enzyme inhibitors functioning at distant points in the metabolic network), as revealed by metabolic control analysis (Kell and Westerhoff, 1986), and thus require separate investigation. The existence of metabolic maps (Baxevanis *et al*. 2002; Kanehisa *et al*. 2002) allows for rapid visualization of metabolite changes and determination of the biological impact of the studied conditions on the organism (e.g. which pathways are affected in drug resistant strains).

Research at the metabolic level is particularly relevant for parasite biology, where metabolic processes are among the major drug targets (for example: dihydrofolate reductase (Methotrexate), ergosterol (Amphotericin B), sterol biosynthesis (azoles), phospholipid metabolism (Miltefosine), and polyamine biosynthesis (Eflornithine)). More generally, it was found that profound metabolic disturbances underlie the lethal effects of almost all drugs. Metabolomics will not only allow for a better understanding of the basic biological processes already partly addressed by transcriptome and proteome studies, but can also be applied to identify metabolic signatures associated with pharmaceutical interventions. Metabolomics allows both (1) profiling the metabolite dynamics associated with the

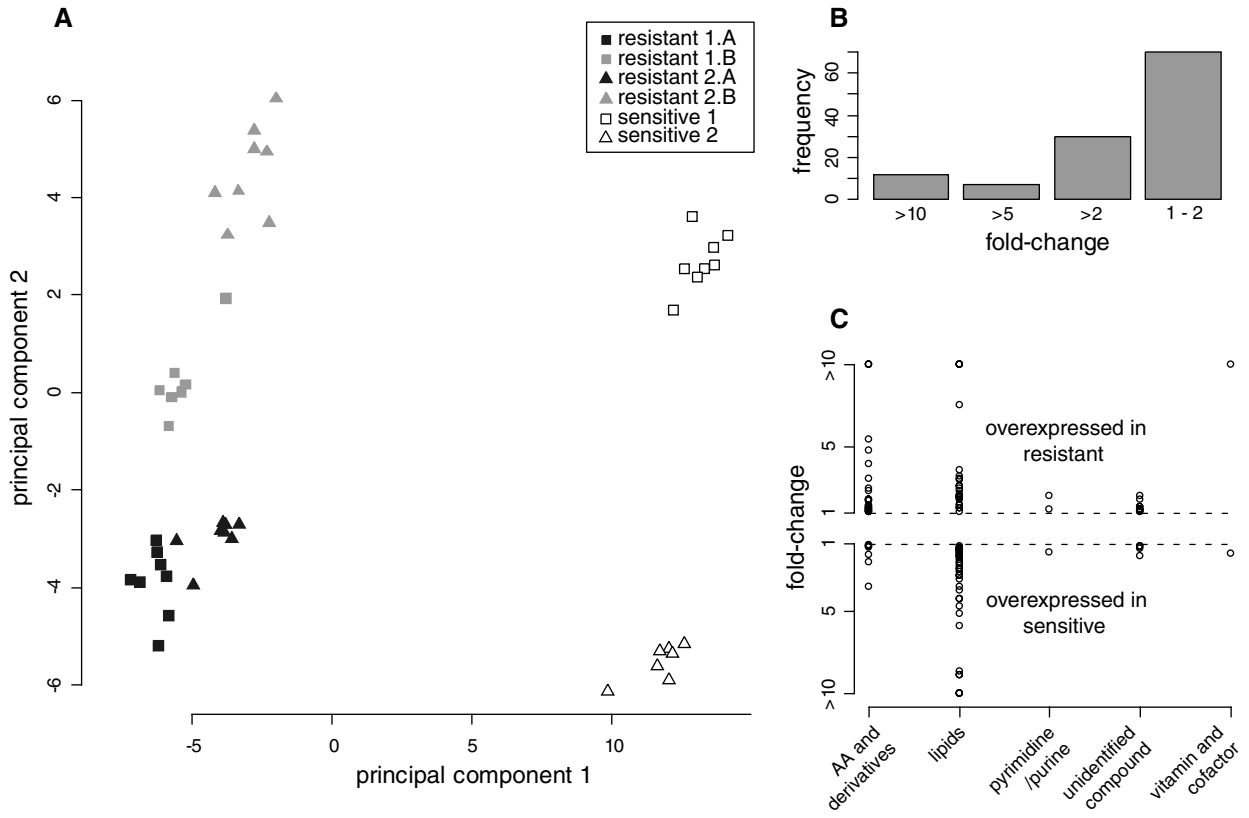

Fig. 2. Principal component analysis can be used to visualize the data so that the measured variation between samples is maximally visible; here it successfully distinguishes drug-sensitive and drug-resistant samples. (A) Each set of biological and technical replicates is clustered closely together, indicating that cell cultures were reproducibly generated and extracted. Principal component 1 clearly separates the two phenotypes and explains 54·4% of the total variance, while principal component 2 separates the different experiments (samples 1 *vs* samples 2) and explains 12·1% of the total variance. (B) The distribution of fold-changes between sensitive and resistant strains shows that a number of metabolites exhibit large, almost binary (>10-fold) differences between the two phenotypes. A total of 49 (41·1%) metabolites show difference of more than 2-fold. (C) Zooming in on the differences per metabolite class shows variability of the relative expression in sensitive *vs* resistant strains across different metabolite classes.

(often unknown) mode of action of effective antiparasitic drugs, and (2) profiling the adapted metabolism of drug-resistant parasites. Metabolic biomarkers could ultimately be reliable predictors of treatment outcome and related clinical polymorphism, in particular because the concentration of metabolites integrates changes happening at both the genomic and environmental level. The effectiveness of metabolomics for signature identification is illustrated in Fig. 2A, which depicts the results of an untargeted metabolomics comparison of drug-resistant and drug-sensitive strains, using metabolite extracts from *L. donovani* cultures measured on an LTQ Orbitrap mass spectrometer combined with HILIC liquid chromatography (Decuypere *et al.* unpublished). Resistant and sensitive strains are clearly separated on the first principal component (explaining 54·4% of the total variance), as over 40% of the identified metabolites show changes of larger than two-fold between the two phenotypes (Fig. 2B). A large portion of these changes is located on well understood pathways involving amino acids (and derivatives) and lipids (Fig. 2C), suggesting hypotheses on the biological impact of the change towards drug resistance. The difference between strains is also clearly evident in a heatmap representation based on an unsupervised hierarchical clustering of the same metabolomics data (Fig. 3).

SAMPLE PREPARATION

Production of reproducible cell cultures and reliable, quantitative extraction of metabolites is the most vital step of any metabolomics study. It is also very challenging, as, for example, the presence of contamination or other confounding factors during sample preparation will likely influence the outcome of the experiment. The lack of well-designed experimental metabolomics reference studies on parasites makes this task even more challenging. The close causal relation between the metabolome and the phenotype makes the metabolome susceptible to any changes in growth conditions (medium, temperature, growth rate, etc.), making it difficult to

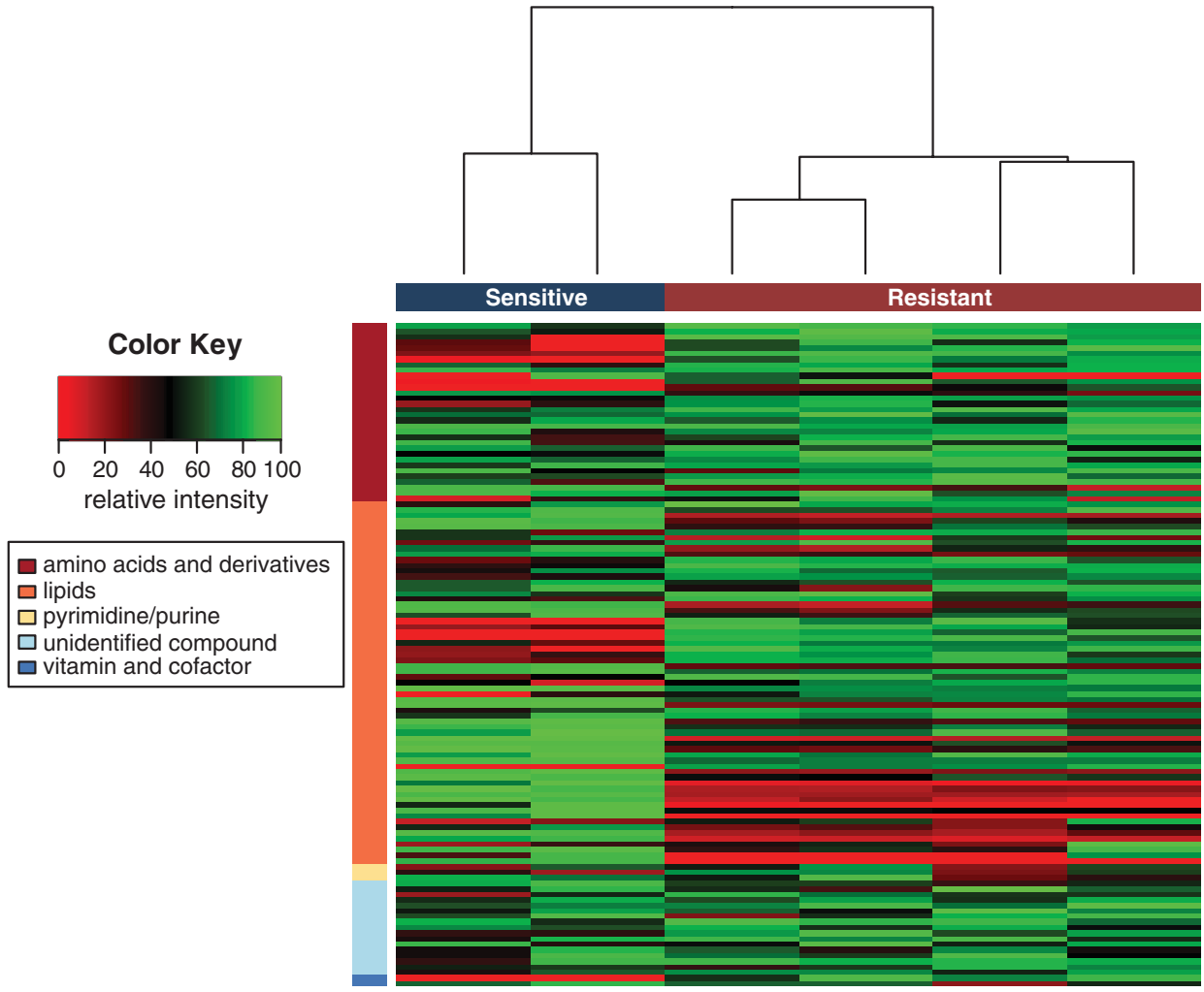

Fig. 3. Hierarchical clustering shows clear separation of the metabolite profiles of sensitive and resistant strains. As an alternative to PCA, the unsupervised hierarchical clustering method can be used to visualise the metabolite expression levels. The similarity of samples is assessed by calculating the correlation between their metabolite profiles (detected intensity levels), and samples that are more similar are closer in the tree. This method shows that large differences between sensitive and resistant strains are observed for almost all the metabolite classes.

obtain reproducible cell cultures. When comparing several strains on a global metabolite level at one particular growth stage, it is an absolute necessity to collect samples from parasite cultures perfectly synchronised to be at the same stage of growth. In order to monitor the events at the metabolic level in more detail, each study should include time course based comparisons. This approach will ultimately also help to visualise metabolite dynamics.

In general, metabolite extracts used for metabolomics should contain as many different metabolites as possible, of different chemical classes, in their original state, and in a quantitative and non-biased manner (Dunn *et al*. 2005; Lu *et al*. 2008). Quenching and cell permeabilization are necessities for achieving this, as they guarantee an accurate snapshot of the physiological conditions of intracellular metabolites. The quenching procedure effectively stops metabolism in all the cells in the culture and needs to act rapidly, as the turnover of biologically important metabolites, such as ATP or NADH, can be in the order of fluxes of millimoles per litre per second in living cells (Winder *et al*. 2008). Great care also needs to be exercised as seemingly innocuous steps such as centrifugation or washing the cells can lead to drastic alterations in intracellular metabolite levels. Quenching is usually done by rapidly changing the cell temperature or pH. For single-cell species such as *Leishmania* several methods exist, including: rapid filtration through an ultrafiltration membrane and subsequent flash-freezing in liquid nitrogen (Wittmann *et al*. 2004), or, more frequently, dilution of cells in cold organic solvent (e.g. methanol, glycerol), with or without buffer (Faijes *et al*. 2007; Villas-Bôas and Bruheim, 2007; Sellick *et al*. 2009). Due to chemical and physical dissimilarities between the cell envelopes of single cell organisms, these approaches do not work equally well for all organisms. Leakage of intracellular metabolites, for example, is practically unavoidable (Winder *et al*.

2008) and can have a strong impact on the observed metabolite concentrations in the extract.

Due to the vast chemical variability of metabolites, covering the complete metabolome in a single experiment is practically impossible. Procedures for non-targeted metabolite extraction focus mainly on obtaining the largest possible representation of relevant metabolites. As parasite cells tend to easily burst when organic solvents are added, hot ethanol extraction (Kamleh *et al*. 2008), and chloroform/methanol/water extraction (De Souza *et al*. 2006) are the current methods of choice for untargeted metabolomics in protozoan parasites. When lipids are targeted, a Bligh–Dyer extraction is recommended as the best approach (Gibellini *et al*. 2009).

METABOLOME ANALYSIS

By far the most popular platform for quantification of analytes at a concentration ranging from subpicomolar to molar levels is mass spectrometry in combination with liquid chromatography (LC–MS), which enables rapid detection of a variety of soluble analytes in a complex sample (Dunn *et al*. 2005). The combination with liquid chromatography reduces the complexity of the mass spectrum at each time along the chromatographic run, potentially separates isomers, and reduces undesired side-effects such as ion suppression (the presence of one compound strongly affecting the signal of another metabolite). It also provides a characteristic retention time for each metabolite, which assists in distinguishing the compound and determining its general chemical properties. The increase in sensitivity and resolution of LC–MS achieved in recent years, most notably with the introduction of the Orbitrap (Hardman and Makarov, 2003), allows for the reliable detection of even those metabolites present at small concentrations in a diverse and variable biological matrix. Continuing improvements of mass spectrometry technology will keep pushing the limits of detectability even further (Han *et al*. 2008).

LC–MS has already been successfully applied to a wide variety of organisms, most notably plants and bacteria (Fiehn, 2001; Kell, 2004). A review of recently published results indicates that a single experiment can quantify between 100 and 200 metabolites in whole-cell extracts (Atherton *et al*. 2006; Dunn *et al*. 2008), corresponding to roughly 10% of the predicted core metabolism. Such a figure stresses the need for more comprehensive techniques to generate an overview of the full complexity of the metabolome. For example, a more satisfactory coverage with LC–MS can be achieved by combining measurements from different chromatographic separation techniques (reversed phase chromatography, hydrophilic interaction chromatography, and ion pair chromatography). Additionally, the current use of LC–MS does not allow for quantification of all detected metabolites (although quantification of selected compounds is possible using C13-labelled standards). This is partly solved by additional measurements using nuclear magnetic resonance spectrometry, which offers the ability to quantify highly abundant metabolites. In Table 1 (Atherton *et al*. 2006; van der Werf *et al*. 2007) a current overview of the advantages and disadvantages of commonly used analytical techniques is provided, including an estimated number of identifiable metabolites and their applicability or popularity within the field of metabolomics.

BIOINFORMATICS CHALLENGES

The sensitivity and resolution currently achieved in mass spectrometry enables the use of complex samples in elaborate experiments (e.g. collecting samples along finely resolved time-series). These can potentially yield large informative datasets about parasite physiology very rapidly, but their interpretation will require extensive bioinformatics efforts. Compared with gene expression microarrays, metabolite mass spectra are far more complex and impossible to interpret manually on a large scale. For example, deconvolution of the detected signals is not yet an established routine operation.

Data generated by mass spectrometry platforms contain large amounts of signals, few of which can be identified by database matching (Fig. 4A). Many of the unidentified signals correspond to analyte derivatives (e.g. isotope peaks, adducts, fragments, multiply charged molecules) and are of interest for the identification of signals, but not for the biological interpretation of the measurements. In order to ease identification, reduce false positives and locate novel analytes, these signals need to be tagged in the data. Such procedures are already well-established for GC-MS data, but are lacking for LC-MS data. The combination of LC–MS with complex experiments such as high-resolution time series or multiple perturbations helps achieving this aim (Tautenhahn *et al*. 2007). Using data-driven clustering based on signal shape correlation and intensity patterns across biological replicates, derivative signals can be identified reliably. In our experiments on *Leishmania* parasites we have found that over 60% of the detected signals consist of derivatives (Decuypere *et al*. unpublished). Time series not only allow for determination of metabolome dynamics following drug treatment or during cellular differentiation, but the intensity correlation pattern among adjacent time points can also be used for sensitive detection of measurement artefacts or spurious fluctuations in metabolites levels (Fig. 4B).

Signals can be putatively identified by matching them on mass to species-specific metabolite databases such as LeishCyc (Doyle *et al*. 2009), compound class-specific databases such as LipidMAPS

Table 1. Each analytical technique suitable for metabolite profiling has specific strengths and weaknesses. Generally, a single analytical method achieves a coverage of roughly 5–10% of the expected metabolome, which can be extended by use of a combination of complementary analytical techniques (van der Werf *et al*. 2007). The number presented in the column 'metabolites measured' provides a rough estimate of the number of metabolites that can be identified using the described analytical method (*, using a single chromatographic column).

Analytical platform	Major advantages	Major disadvantages	Metabolites measured	Applications
'*Single' detection techniques*				
Nuclear Magnetic Resonance (NMR)	• detection independent of ionisation • structural characterization power	• long acquisition time • low resolution and sensitivity	20–50	• fast screening method for abundant metabolites • *de novo* identification
Mass Spectrometry (MS)	• fast and sensitive detection of metabolite masses	• requires ions for detection • competitive ionisation of metabolites in complex mixtures leads to ionization suppression • inability to discriminate between isomers • adduct formation	n.a.	
Hyphenated analytical techniques				
Gas Chromatography (GC)–MS (*separation based on volatility*)	• good separation efficiency • low-cost and easy-to-operate system • excellent deconvolution and metabolite identification software • derivatisation increases number of detectable metabolites	• metabolites need to be volatile and thermostable, difficult for higher mass molecules or molecules not easily rendered volatile • derivatization includes more sample handling, adds up in time and generates more analytical variance	200–300	• gold standard in metabolomics analysis
Liquid Chromatography (LC)–MS (*separation based on chemical interaction*)	• excellent versatility due to broad range of columns with different chemical properties • good separation efficiency • good dynamic range • potentially separates isomers	• complex data	100–200*	• most used platform in metabolomics analysis
Capillary Electrophoresis (CE)–MS (*separation based on mass-to-charge ratio*)	• versatile and fast method • good separation efficiency • low running costs	• barely used in comprehensive metabolomics analysis	50–150	• mainly complementary metabolomics technique, i.e. when a major fraction of the sample consists of polar and ionic metabolites • upcoming metabolomics platform

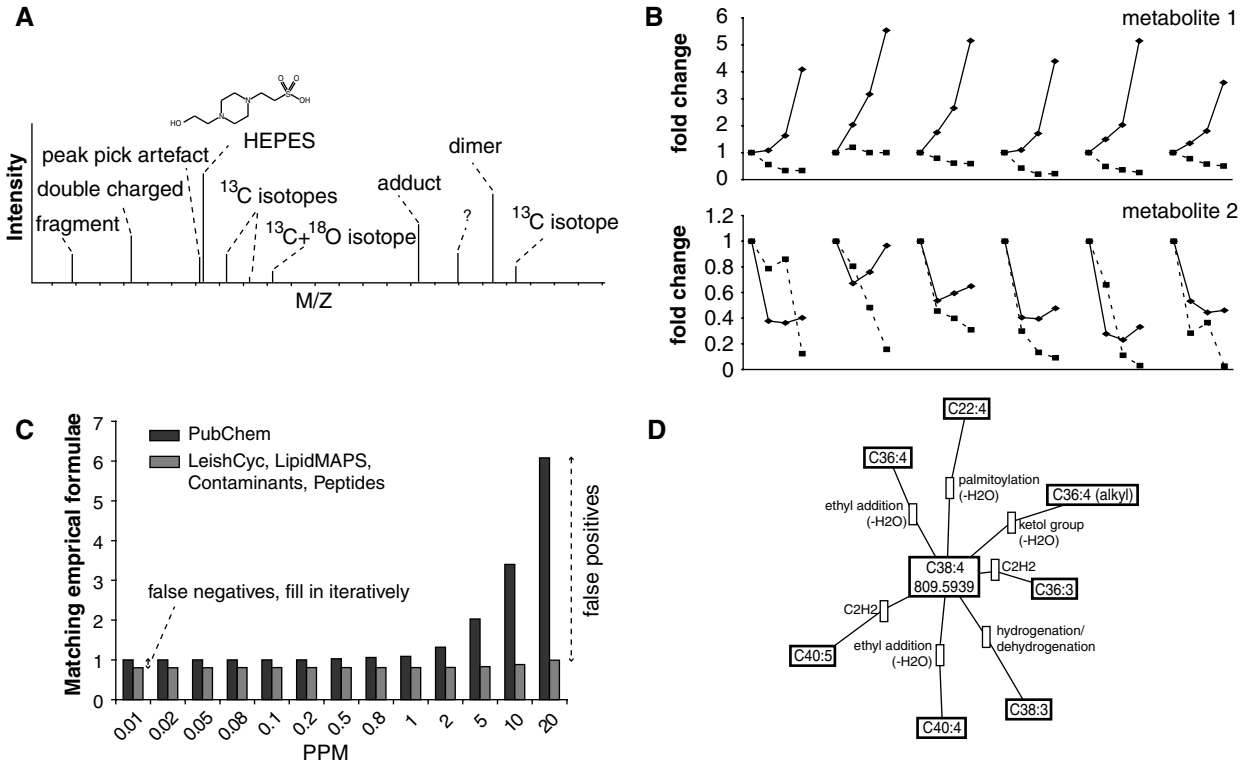

Fig. 4. Illustrations of some bioinformatics challenges for metabolomics data. (A) Mass spectrometry data contain a large amount of signals, most of which can be identified as derivatives of a metabolite of interest (e.g. isotopes, adducts). Data-driven reduction approaches help to reveal these derivative signals and reduce the annotation burden. (B) Time series data can reveal metabolite dynamics following drug treatment or during differentiation, but also provide important validation of observed expression levels. Here 6 biological replicates with (dashed) and without (solid) perturbation are tracked over 4 time points for two metabolites, providing insight into their dynamic response to perturbation. (C) After the applied deconvolution approaches, the detected compounds need to be identified. The graph shows how the use of specific databases (~10·000 compounds; light-grey bars), in combination with higher mass accuracy, results in a drastically lower number of false positives compared to using the comprehensive PubChem database (more than 40 million compounds; dark-grey bars),; this can significantly reduce the burden of manual curation. Interestingly, PubChem will always result in a match, while the collection of databases does not (false negatives; number of matches smaller than 1). To fill in the gaps, the remaining unidentified signals are iteratively assigned putative identities using progressively less specific databases. The resulting putative identifications then need to be verified with orthogonal identification techniques (e.g. fragmentation studies). (D) Using the observed exact mass differences between metabolites it is possible to organise the observations in hypothetical metabolic networks by linking them with a discrete set of biochemical transformations, further facilitating the interpretation process.

(Fahy *et al*. 2007), and contaminant databases, the latter allows removal of typical impurities and buffer components often detected in metabolomics experiments (Keller *et al*. 2008). For many samples it is also useful to specifically search for computer-generated lists of, for example, possible di- and tripeptides. Together, this set of databases is a best first guess of the identity of the expected metabolites for a *Leishmania* sample, and results in most cases in a single chemical formula per signal (Fig. 4C). Remaining unidentified compounds can then be matched iteratively to more comprehensive databases, including KEGG (Kanehisa *et al*. 2002), Metlin (Smith *et al*. 2005), the Human Metabolome Database (Wishart *et al*. 2007) or PubChem (Wang *et al*. 2009), in roughly this order. These larger databases tend to give more spurious hits, and careful manual curation of the results is required. The concurrent use of gas chromatography coupled to mass spectrometry (GC–MS) in a comprehensive platform will greatly enhance the unambiguous identification of metabolites. The stable retention times produced with this method and the existence of libraries containing retention time information for a huge array of analytes allows almost complete automation of this process (Pasikanti *et al*. 2008).

The high mass accuracy has additional advantages. An area of particular interest for parasite metabolomics is the *ab initio* extension of the metabolic network (Fig. 4D): based on the extremely accurate mass information provided by the newest mass spectrometers, it is possible to identify potential metabolic transformations connecting the observed mass peaks (Breitling *et al*. 2006). These hypothetical connections between metabolites can be used to explore the metabolic context of unidentified

compounds, to highlight connected areas of the metabolome that show concerted changes in response to a perturbation, such as drug treatment or cellular differentiation, and to expand the computational genome-based reconstructions of the metabolic network to include novel pathways that may be species-specific (Breitling *et al*. 2008). This may be particularly relevant for *Leishmania*: after all, roughly 1000 genes are specific to this genus, and only very few of them have been functionally characterized (Peacock *et al*. 2007). The network context can also be used to support identifications of metabolites (Rogers *et al*. 2009), which is further supported by the combination with tandem mass spectrometry fragmentation patterns (Bocker *et al*. 2008).

FUTURE PERSPECTIVE

Once metabolomics has been established as a routine tool for *Leishmania* research, the next challenge will be the integration of data across the various molecular levels. A comprehensive understanding of *Leishmania* diversity is likely to require work at the genome, proteome and metabolome levels in a unified framework. With the current technological advances, such a framework has come within reach. The new genome sequencing technologies developed over the past years, allow whole genome comparative analyses of multiple strains of a given species. We expect that the integration of global genomic and metabolomic properties will result in unprecedented insight into parasite diversity at the molecular level. The exploitation of the resulting systems-wide view and its effective translation to treatment strategies will require new alliances between specialized high-tech centres, developing advanced measurement technologies, and institutions engaged in the field, at the frontline of infectious disease research. A first initiative for post-genomic integration in *Leishmania* research is the GeMInI consortium (http://www.leishrisk.net/gemini), which brings together next generation sequencing and high-performance metabolomics on drug-sensitive and resistant *Leishmania donovani* clinical isolates.

ACKNOWLEDGEMENTS

The authors thank the British Society for Parasitology for organizing the Autumn Symposium 2009 and offering the opportunity to publish this review. We also thank Dave Watson (Strathclyde University) for the fruitful collaboration on generating high-quality metabolomics data for *Leishmania*, and Marnix Medema and Andris Jankevics for their helpful comments.

FINANCIAL SUPPORT

This work was supported by an NWO-Vidi grant to RB. The work of RAS was funded by an NWO-Vici grant to Ritsert Jansen (Groningen). The GeMInI initiative of the Institute of Tropical Medicine, a Marie Curie Intra-European Fellowship (contract DonNA-R – 041465) and the Baillet–Latour Foundation provided financial support for SD.

REFERENCES

Atherton, H. J., Bailey, N. J., Zhang, W., Taylor, J., Major, H., Shockcor, J., Clarke, K. and Griffin, J. L. (2006). A combined 1H-NMR spectroscopy- and mass spectrometry-based metabolomic study of the PPAR-alpha null mutant mouse defines profound systemic changes in metabolism linked to the metabolic syndrome. *Physiological Genomics* **27**, 178–186. doi: 10.1152/physiolgenomics.00060.2006.

Baxevanis, A. D., Page, R. D., Petsko, G. A., Stein, L. D. and Stormo, G. D. (2002). *Current Protocols in Bioinformatics*. John Wiley & Sons, Inc., Hoboken, NJ, USA.

Bente, M., Harder, S., Wiesgigl, M., Heukeshoven, J., Gelhaus, C., Krause, E., Clos, J. and Bruchhaus, I. (2003). Developmentally induced changes of the proteome in the protozoan parasite *Leishmania donovani*. *Proteomics* **3**, 1811–1829. doi: 10.1002/pmic.200300462.

Bocker, S. and Rasche, F. (2008). Towards *de novo* identification of metabolites by analyzing tandem mass spectra. *Bioinformatics* **24**, i49–55. doi: 10.1093/bioinformatics/btn270.

Breitling, R., Ritchie, S., Goodenowe, D., Stewart, M. L. and Barrett, M. P. (2006). *Ab initio* prediction of metabolic networks using Fourier transform mass spectrometry data. *Metabolomics* **2**, 155–164. doi: 10.1007/s11306-006-0029-z.

Breitling, R., Vitkup, D. and Barrett, M. P. (2008). New surveyor tools for charting microbial metabolic maps. *Nature Reviews, Microbiology* **6**, 156–161. doi: 10.1038/nrmicro1797.

Brobey, R. K. B., Mei, F. C., Cheng, X. and Soong, L. (2006). Comparative two-dimensional gel electrophoresis maps for promastigotes of *Leishmania amazonensis* and *Leishmania major*. *Brazilian Journal of Infectious Diseases : An Official Publication of the Brazilian Society of Infectious Diseases* **10**, 1–6. doi: /S1413-86702006000100001.

Chavali, A. K., Whittemore, J. D., Eddy, J. A., Williams, K. T. and Papin, J. A. (2008). Systems analysis of metabolism in the pathogenic trypanosomatid *Leishmania major*. *Molecular Systems Biology* **4**, 177. doi: 10.1038/msb.2008.15.

Clayton, C. and Shapira, M. (2007). Post-transcriptional regulation of gene expression in trypanosomes and leishmanias. *Molecular and Biochemical Parasitology* **156**, 93–101. doi: 10.1016/j.molbiopara.2007.07.007.

Cohen-Freue, G., Holzer, T. R., Forney, J. D. and McMaster, W. R. (2007). Global gene expression in *Leishmania*. *International Journal for Parasitology* **37**, 1077–1086. doi: 10.1016/j.ijpara.2007.04.011.

De Souza, D. P., Saunders, E. C., McConville, M. J. and Likic, V. A. (2006). Progressive peak clustering in GC-MS Metabolomic experiments applied to *Leishmania* parasites. *Bioinformatics* **22**, 1391–1396. doi: 10.1093/bioinformatics/btl085.

Doyle, M., MacRae, J., De Souza, D., Saunders, E., McConville, M. and Likic, V. (2009). LeishCyc: a biochemical pathways database for *Leishmania major*. *BMC Systems Biology* **3**, 57. doi: 10.1186/1752-0509-3-57.

Drummelsmith, J., Brochu, V., Girard, I., Messier, N. and Ouellette, M. (2003). Proteome mapping of the protozoan parasite *Leishmania* and application to the study of drug targets and resistance mechanisms. *Molecular and Cellular Proteomics* **2**, 146–155. doi: 10.1074/mcp.M200085-MCP200.

Drummelsmith, J., Girard, I., Trudel, N. and Ouellette, M. (2004). Differential protein expression analysis of *Leishmania major* reveals novel roles for methionine adenosyltransferase and S-adenosylmethionine in methotrexate resistance. *Journal of Biological Chemistry* **279**, 33273–33280. doi: 10.1074/jbc.M405183200.

Dujardin, J. (2009). Structure, dynamics and function of *Leishmania* genome: resolving the puzzle of infection, genetics and evolution? *Infection, Genetics and Evolution: Journal of Molecular Epidemiology and Evolutionary Genetics in Infectious Diseases* **9**, 290–297. doi: 10.1016/j.meegid.2008.11.007.

Dunn, W. B., Bailey, N. J. C. and Johnson, H. E. (2005). Measuring the metabolome: current analytical technologies. *The Analyst* **130**, 606–625.

Dunn, W. B., Broadhurst, D., Brown, M., Baker, P. N., Redman, C. W. G., Kenny, L. C. and Kell, D. B. (2008). Metabolic profiling of serum using Ultra Performance Liquid Chromatography and the LTQ-Orbitrap mass spectrometry system. *Journal of Chromatography B* **871**, 288–298. doi: 10.1016/j.jchromb.2008.03.021.

Fadili, K. E., Drummelsmith, J., Roy, G., Jardim, A. and Ouellette, M. (2009). Down regulation of KMP-11 in *Leishmania infantum* axenic antimony resistant amastigotes as revealed by a proteomic screen. *Experimental Parasitology* **123**, 51–57. doi: 10.1016/j.exppara.2009.05.013.

Fahy, E., Sud, M., Cotter, D. and Subramaniam, S. (2007). LIPID MAPS online tools for lipid research. *Nucleic Acids Research* **35**(Web Server issue), W606–612. doi: 10.1093/nar/gkm324.

Faijes, M., Mars, A. and Smid, E. (2007). Comparison of quenching and extraction methodologies for metabolome analysis of *Lactobacillus plantarum*. *Microbial Cell Factories* **6**, 27. doi: 10.1186/1475-2859-6-27.

Fernie, A. R., Trethewey, R. N., Krotzky, A. J. and Willmitzer, L. (2004). Metabolite profiling: from diagnostics to systems biology. *Nature Reviews Molecular Cell Biology* **5**, 763–769. doi: 10.1038/nrm1451.

Fiehn, O. (2001). Combining genomics, metabolome analysis, and biochemical modelling to understand metabolic networks. *Comparative and Functional Genomics* **2**, 155–168. doi: 10.1002/cfg.82.

Gibellini, F., Hunter, W. N. and Smith, T. K. (2009). The ethanolamine branch of the Kennedy pathway is essential in the bloodstream form of *Trypanosoma brucei*. *Molecular Microbiology* **73**, 826–843. doi: 10.1111/j.1365-2958.2009.06764.x.

Goodacre, R., Vaidyanathan, S., Dunn, W. B., Harrigan, G. G. and Kell, D. B. (2004). Metabolomics by numbers: acquiring and understanding global metabolite data. *Trends in Biotechnology* **22**, 245–252. doi: 10.1016/j.tibtech.2004.03.007.

Guerbouj, S., Victoir, K., Guizani, I., Seridi, N., Nuwayri-Salti, N., Belkaid, M., Ismail, R. B., Le Ray, D. and Dujardin, J. C. (2001). Gp63 gene polymorphism and population structure of *Leishmania donovani* complex: influence of the host selection pressure? *Parasitology* **122**, 25–35.

Guimond, C., Trudel, N., Brochu, C., Marquis, N., Fadili, A. E., Peytavi, R., Briand, G., Richard, D., Messier, N., Papadopoulou, B., Corbeil, J., Bergeron, M. G., Legare, D. and Ouellette, M. (2003). Modulation of gene expression in *Leishmania* drug resistant mutants as determined by targeted DNA microarrays. *Nucleic Acids Research* **31**, 5886–5896. doi: 10.1093/nar/gkg806.

Hall, N. (2007). Advanced sequencing technologies and their wider impact in microbiology. *Journal of Experimental Biology* **210**, 1518–1525. doi: 10.1242/jeb.001370.

Han, J., Danell, R., Patel, J., Gumerov, D., Scarlett, C., Speir, J., Parker, C., Rusyn, I., Zeisel, S. and Borchers, C. (2008). Towards high-throughput metabolomics using ultrahigh-field Fourier transform ion cyclotron resonance mass spectrometry. *Metabolomics* **4**, 128–140. doi: 10.1007/s11306-008-0104-8.

Hardman, M. and Makarov, A. A. (2003). Interfacing the orbitrap mass analyzer to an electrospray ion source. *Analytical Chemistry* **75**, 1699–1705.

Holzer, T. R., McMaster, W. and Forney, J. D. (2006). Expression profiling by whole-genome interspecies microarray hybridization reveals differential gene expression in procyclic promastigotes, lesion-derived amastigotes, and axenic amastigotes in *Leishmania mexicana*. *Molecular and Biochemical Parasitology* **146**, 198–218. doi: 10.1016/j.molbiopara.2005.12.009.

Inga, R., Doncker, S. D., Gomez, J., Lopez, M., Garcia, R., Ray, D. L., Arevalo, J. and Dujardin, J. (1998). Relation between variation in copy number of ribosomal RNA encoding genes and size of harbouring chromosomes in *Leishmania* of subgenus *Viannia*. *Molecular and Biochemical Parasitology* **92**, 219–228. doi: 10.1016/S0166-6851(98)00009-7.

Ivens, A. C., Peacock, C. S., Worthey, E. A., Murphy, L., Aggarwal, G., Berriman, M., Sisk, E., Rajandream, M., Adlem, E., Aert, R., Anupama, A., Apostolou, Z., Attipoe, P., Bason, N., Bauser, C., Beck, A., Beverley, S. M., Bianchettin, G., Borzym, K., Bothe, G., Bruschi, C. V., Collins, M., Cadag, E., Ciarloni, L., Clayton, C., Coulson, R. M. R., Cronin, A., Cruz, A. K., Davies, R. M., De Gaudenzi, J., Dobson, D. E., Duesterhoeft, A., Fazelina, G., Fosker, N., Frasch, A. C., Fraser, A., Fuchs, M., Gabel, C., Goble, A., Goffeau, A., Harris, D., Hertz-Fowler, C., Hilbert, H., Horn, D., Huang, Y., Klages, S., Knights, A., Kube, M., Larke, N., Litvin, L., Lord, A., Louie, T., Marra, M., Masuy, D., Matthews, K., Michaeli, S., Mottram, J. C., Müller-Auer, S., Munden, H., Nelson, S., Norbertczak, H., Oliver, K., O'Neil, S., Pentony,

M., Pohl, T. M., Price, C., Purnelle, B., Quail, M. A., Rabbinowitsch, E., Reinhardt, R., Rieger, M., Rinta, J., Robben, J., Robertson, L., Ruiz, J. C., Rutter, S., Saunders, D., Schäfer, M., Schein, J., Schwartz, D. C., Seeger, K., Seyler, A., Sharp, S., Shin, H., Sivam, D., Squares, R., Squares, S., Tosato, V., Vogt, C., Volckaert, G., Wambutt, R., Warren, T., Wedler, H., Woodward, J., Zhou, S., Zimmermann, W., Smith, D. F., Blackwell, J. M., Stuart, K. D., Barrell, B. and Myler, P. J. (2005). The genome of the kinetoplastid parasite, *Leishmania major*. *Science* **309**, 436–442. doi: 10.1126/science.1112680.

Kamleh, A., Barrett, M. P., Wildridge, D., Burchmore, R. J. S., Scheltema, R. A. and Watson, D. G. (2008). Metabolomic profiling using Orbitrap Fourier transform mass spectrometry with hydrophilic interaction chromatography: a method with wide applicability to analysis of biomolecules. *Rapid Communications in Mass Spectrometry* **22**, 1912–1918. doi: 10.1002/rcm.3564.

Kanehisa, M., Goto, S., Kawashima, S. and Nakaya, A. (2002). The KEGG databases at GenomeNet. *Nucleic Acids Research* **30**, 42–46.

Kell, D. and Westerhoff, H. (1986). Metabolic control theory: its role in microbiology and biotechnology. *FEMS Microbiology Letters* **39**, 305–320. doi: 10.1111/j.1574-6968.1986.tb01863.x.

Kell, D. B. (2004). Metabolomics and systems biology: making sense of the soup. *Current Opinion in Microbiology* **7**, 296–307. doi: 10.1016/j.mib.2004.04.012.

Keller, B. O., Sui, J., Young, A. B. and Whittal, R. M. (2008). Interferences and contaminants encountered in modern mass spectrometry. *Analytica Chimica Acta* **627**, 71–81. doi: 10.1016/j.aca.2008.04.043.

Leifso, K., Cohen-Freue, G., Dogra, N., Murray, A. and McMaster, W. R. (2007). Genomic and proteomic expression analysis of *Leishmania* promastigote and amastigote life stages: The *Leishmania* genome is constitutively expressed. *Molecular and Biochemical Parasitology* **152**, 35–46. doi: 10.1016/j.molbiopara.2006.11.009.

Leprohon, P., Legare, D., Raymond, F., Madore, E., Hardiman, G., Corbeil, J. and Ouellette, M. (2009). Gene expression modulation is associated with gene amplification, supernumerary chromosomes and chromosome loss in antimony-resistant *Leishmania infantum*. *Nucleic Acids Research* **37**, 1387–1399. doi: 10.1093/nar/gkn1069.

Lu, X., Zhao, X., Bai, C., Zhao, C., Lu, G. and Xu, G. (2008). LC-MS-based metabonomics analysis. *Journal of Chromatography B* **866**, 64–76. doi: 10.1016/j.jchromb.2007.10.022.

McConville, M. J., de Souza, D., Saunders, E., Likic, V. A. and Naderer, T. (2007). Living in a phagolysosome; metabolism of *Leishmania* amastigotes. *Trends in Parasitology* **23**, 368–375. doi: 10.1016/j.pt.2007.06.009.

McDonagh, P. D., Myler, P. J. and Stuart, K. (2000). The unusual gene organization of *Leishmania major* chromosome 1 may reflect novel transcription processes. *Nucleic Acids Research* **28**, 2800–2803. doi: 10.1093/nar/28.14.2800.

McNicoll, F., Drummelsmith, J., Müller, M., Madore, É., Boilard, N., Ouellette, M. and Papadopoulou, B. (2006). A combined proteomic and transcriptomic approach to the study of stage differentiation in *Leishmania infantum*. *Proteomics* **6**, 3567–3581. doi: 10.1002/pmic.200500853.

Morozova, O. and Marra, M. A. (2008). Applications of next-generation sequencing technologies in functional genomics. *Genomics* **92**, 255–264. doi: 10.1016/j.ygeno.2008.07.001.

Ouellette, M., Ubeda, J., Leprohon, P., Mukherjee, A., Brotherton, C., Moreira, W., Coelho, A. and Raymond, F. (2009). Whole genome analysis of drug resistance in *Leishmania*. In *Abstracts* World Conference Leishmaniasis 4.

Pasikanti, K. K., Ho, P. and Chan, E. (2008). Gas chromatography/mass spectrometry in metabolic profiling of biological fluids. *Journal of Chromatography B* **871**, 202–211. doi: 10.1016/j.jchromb.2008.04.033.

Peacock, C. S., Seeger, K., Harris, D., Murphy, L., Ruiz, J. C., Quail, M. A., Peters, N., Adlem, E., Tivey, A., Aslett, M., Kerhornou, A., Ivens, A., Fraser, A., Rajandream, M., Carver, T., Norbertczak, H., Chillingworth, T., Hance, Z., Jagels, K., Moule, S., Ormond, D., Rutter, S., Squares, R., Whitehead, S., Rabbinowitsch, E., Arrowsmith, C., White, B., Thurston, S., Bringaud, F., Baldauf, S. L., Faulconbridge, A., Jeffares, D., Depledge, D. P., Oyola, S. O., Hilley, J. D., Brito, L. O., Tosi, L. R. O., Barrell, B., Cruz, A. K., Mottram, J. C., Smith, D. F. and Berriman, M. (2007). Comparative genomic analysis of three *Leishmania* species that cause diverse human disease. *Nature Genetics* **39**, 839–847. doi: 10.1038/ng2053.

Rochette, A., Raymond, F., Ubeda, J., Smith, M., Messier, N., Boisvert, S., Rigault, P., Corbeil, J., Ouellette, M. and Papadopoulou, B. (2008). Genome-wide gene expression profiling analysis of *Leishmania major* and *Leishmania infantum* developmental stages reveals substantial differences between the two species. *BMC Genomics* **9**, 255. doi: 10.1186/1471-2164-9-255.

Rogers, S., Scheltema, R. A., Girolami, M. and Breitling, R. (2009). Probabilistic assignment of formulas to mass peaks in metabolomics experiments. *Bioinformatics* **25**, 512–518. doi: 10.1093/bioinformatics/btn642.

Rosenzweig, D., Smith, D., Opperdoes, F., Stern, S., Olafson, R. W. and Zilberstein, D. (2008). Retooling *Leishmania* metabolism: from sand fly gut to human macrophage. *FASEB Journal* **22**, 590–602. doi: 10.1096/fj.07-9254com.

Salotra, P., Duncan, R. C., Singh, R., Raju, B. S., Sreenivas, G. and Nakhasi, H. L. (2006). Upregulation of surface proteins in *Leishmania donovani* isolated from patients of post kala-azar dermal leishmaniasis. *Microbes and Infection* **8**, 637–644. doi: 10.1016/j.micinf.2005.08.018.

Saxena, A., Lahav, T., Holl, N., Aggarwal, G., Anupama, A., Huang, Y., Volpin, H., Myler, P. and Zilberstein, D. (2007). Analysis of the *Leishmania donovani* transcriptome reveals an ordered progression of transient and permanent changes in gene expression during differentiation. *Molecular*

and Biochemical Parasitology **152**, 53–65. doi: 10.1016/j.molbiopara.2006.11.011.

Sellick, C. A., Hansen, R., Maqsood, A. R., Dunn, W. B., Stephens, G. M., Goodacre, R. and Dickson, A. J. (2009). Effective quenching processes for physiologically valid metabolite profiling of suspension cultured mammalian cells. *Analytical Chemistry* **81**, 174–183. doi: 10.1021/ac8016899.

Sharma, S., Singh, G., Chavan, H. D. and Dey, C. S. (2003). Proteomic analysis of wild type and arsenite-resistant *Leishmania donovani*. *Experimental Parasitology* **123**, 369–376. doi: 10.1016/j.exppara.2009.08.003.

Singh, G., Chavan, H. D. and Dey, C. S. (2008). Proteomic analysis of miltefosine-resistant *Leishmania* reveals the possible involvement of eukaryotic initiation factor 4A (eIF4A). *International Journal of Antimicrobial Agents* **31**, 584–586. doi: 10.1016/j.ijantimicag.2008.01.032.

Singh, N., Almeida, R., Kothari, H., Kumar, P., Mandal, G., Chatterjee, M., Venkatachalam, S., Govind, M. K., Mandal, S. K. and Sundar, S. (2007). Differential gene expression analysis in antimony-unresponsive Indian kala azar (visceral leishmaniasis) clinical isolates by DNA microarray. *Parasitology* **134**, 777–787. doi: 10.1017/S0031182007002284.

Smith, C. A., O'Maille, G., Want, E. J., Qin, C., Trauger, S. A., Brandon, T. R., Custodio, D. E., Abagyan, R. and Siuzdak, G. (2005). METLIN: a metabolite mass spectral database. *Therapeutic Drug Monitoring* **27**, 747–751.

Smith, D. F., Peacock, C. S. and Cruz, A. K. (2007). Comparative genomics: from genotype to disease phenotype in the leishmaniases. *International Journal for Parasitology* **37**, 1173–1186. doi: 10.1016/j.ijpara.2007.05.015.

Srividya, G., Duncan, R., Sharma, P., Raju, B. V. S., Nakhasi, H. L. and Salotra, P. (2007). Transcriptome analysis during the process of *in vitro* differentiation of *Leishmania donovani* using genomic microarrays. *Parasitology* **134**, 1527–1539. doi: 10.1017/S003118200700296X.

Sturm, N. R., Martinez, L. I. T. and Thomas, S. (2008). Kinetoplastid genomics: The thin end of the wedge. *Infection, Genetics and Evolution* **8**, 901–906. doi: 10.1016/j.meegid.2008.07.001.

Tautenhahn, R., Böttcher, C. and Neumann, S. (2007). Annotation of LC/ESI-MS Mass Signals. In *Bioinformatics Research and Development*. (ed. Hochreiter, S. and Wagner, R.), pp. 371–380. Springer-Verlag, Heidelberg, Germany.

Ubeda, J., Legare, D., Raymond, F., Ouameur, A., Boisvert, S., Rigault, P., Corbeil, J., Tremblay, M., Olivier, M., Papadopoulou, B. and Ouellette, M. (2008). Modulation of gene expression in drug resistant *Leishmania* is associated with gene amplification, gene deletion and chromosome aneuploidy. *Genome Biology* **9**, R115. doi: 10.1186/gb-2008-9-7-r115.

van der Werf, M. J., Overkamp, K. M., Muilwijk, B., Coulier, L. and Hankemeier, T. (2007). Microbial metabolomics: toward a platform with full metabolome coverage. *Analytical Biochemistry* **370**, 17–25. doi: 10.1016/j.ab.2007.07.022.

Vergnes, B., Gourbal, B., Girard, I., Sundar, S., Drummelsmith, J. and Ouellette, M. (2007). A proteomics screen implicates HSP83 and a small kinetoplastid calpain-related protein in drug resistance in *Leishmania donovani* clinical field isolates by modulating drug-induced programmed cell death. *Molecular and Cellular Proteomics* **6**, 88–101. doi: 10.1074/mcp.M600319-MCP200.

Victoir, K., Dujardin, J. C., de Doncker, S., Barker, D. C., Arevalo, J., Hamers, R. and Le Ray, D. (1995). Plasticity of gp63 gene organization in *Leishmania* (*Viannia*) *braziliensis* and *Leishmania* (*Viannia*) *peruviana*. *Parasitology* **111**, 265–273.

Villas-Bôas, S. G. and Bruheim, P. (2007). Cold glycerol-saline: The promising quenching solution for accurate intracellular metabolite analysis of microbial cells. *Analytical Biochemistry* **370**, 87–97. doi: 10.1016/j.ab.2007.06.028.

Wang, Y., Xiao, J., Suzek, T. O., Zhang, J., Wang, J. and Bryant, S. H. (2009). PubChem: a public information system for analyzing bioactivities of small molecules. *Nucleic Acids Research* **37**, W623–633. doi: 10.1093/nar/gkp456.

Winder, C. L., Dunn, W. B., Schuler, S., Broadhurst, D., Jarvis, R., Stephens, G. M. and Goodacre, R. (2008). Global metabolic profiling of *Escherichia coli* cultures: an evaluation of methods for quenching and extraction of intracellular metabolites. *Analytical Chemistry* **80**, 2939–2948. doi: 10.1021/ac7023409.

Wishart, D. S., Tzur, D., Knox, C., Eisner, R., Guo, A. C., Young, N., Cheng, D., Jewell, K., Arndt, D., Sawhney, S., Fung, C., Nikolai, L., Lewis, M., Coutouly, M., Forsythe, I., Tang, P., Shrivastava, S., Jeroncic, K., Stothard, P., Amegbey, G., Block, D., Hau, D. D., Wagner, J., Miniaci, J., Clements, M., Gebremedhin, M., Guo, N., Zhang, Y., Duggan, G. E., MacInnis, G. D., Weljie, A. M., Dowlatabadi, R., Bamforth, F., Clive, D., Greiner, R., Li, L., Marrie, T., Sykes, B. D., Vogel, H. J. and Querengesser, L. (2007). HMDB: the human metabolome database. *Nucleic Acids Research* **35** (Database issue), D521–526. doi: 10.1093/nar/gkl923.

Wittmann, C., Krömer, J. O., Kiefer, P., Binz, T. and Heinzle, E. (2004). Impact of the cold shock phenomenon on quantification of intracellular metabolites in bacteria. *Analytical Biochemistry* **327**, 135–139. doi: 10.1016/j.ab.2004.01.002.